Integrated Water Quality and Aquatic Communities Protocol – Lakes and Ponds

Standard Operating Procedure (SOP) #17: Data Transfer, Storage, and Archive

Draft Version 1.0

Revision History Log:

Previous Version	Revision Date	Author	Changes Made	Reason for Change	New Version

This SOP explains the procedures for data transfer to the Network Data Manager. In addition, data certification, storage, and archiving and a timeline for project deliverables are addressed.

Data Transfer

All project deliverables, including but not limited to raw data, processed data, Metadata Interview forms, updated data dictionary (if necessary), images with metadata, training logs, datasheets and logs, data entry logs, equipment logs, special event logs, spatial files, and Certification forms will be transferred to the Klamath Network Data Manager following the timeline listed in Table 1. It is the responsibility of the Project Lead to ensure all products and associated documentation are delivered following the outlined timeline. At no point in time should a new field season start if project deliverables have not been completed unless prior permission has been granted by the Network Coordinator and Data Manager.

Certification Form

The Klamath Network will utilize a Certification form submitted by the Project Lead to ensure:

1. The data are complete for the period of time indicated on the form.
2. The data have undergone the quality assurance checks indicated in the SOP#16: Quality Assurance Project Plan.
3. Metadata for all data has been provided (when applicable).
4. Project timelines are being followed and all products from the field season have been submitted.
5. The correct level of sensitivity is associated with the deliverables.

A new Certification form should be submitted each time a product is submitted. If multiple products are submitted at the same time, only one Certification form is needed for those products. Certification Forms can be obtained from the Network Internet and Intranet web sites

or by contacting the Network Data Manager. An example of the Certification form is included at the end of this SOP.

Field Forms

Hardcopies of the datasheets will be provided to the Network Data Manager following the timeline below. It is the responsibility of the Data Manager to scan the datasheets into PDF documents within 1 month of receiving the hardcopies. The datasheets will be organized in the order in which they are to be scanned when they are transferred to the Data Manager (SOP #15: Post-Field Season). For the lakes, there will be three separate sets:

1. Site Field Forms (standard datasheet)
2. Fish Collection Forms
3. Protocol Associated Log Sheets

The details of scanning and naming structure of the files are covered in SOP #15: Post-Field Season.

The scanned log sheets will include:

- Calibration logs
- Training logs
- Crew Expectation logs (Appendix B)
- Chain of Custody Forms
- Any QA/QC documents arising from the QAPP (SOP# 17)

Electronic files will be stored in the documentation folder located at G:\Monitoring\Water_Quality_Monitoring\Lakes\Lakes_Data\Datasheets\Seasonal_Data\YYYY on the Network server. Additional details on storage methods are described below.

Databases

At the end of the field season, the project database will be provided to the Network Data Manager along with the Certification form, Metadata Interview form, and, if necessary, an updated data dictionary (SOP#18: Metadata Guidelines). It is the responsibility of the Project Lead to examine the project database for accuracy and completeness prior to transferring the database to the Data Manager. Once the database(s) have been transferred to the Data Manager, he or she will run the data through one more round of validation/verification checks and then load the data into a master database that contains all the data from previous years.

Photos

Images and associated metadata will be transferred to the Data Manager in the format explained in SOP #14: Photo Management.

GIS Shapefiles

A GIS shapefile is created and data stored therein using ArcPad 8, as described in SOP #11: Amphibian, Invertebrates, and Lake Substrate Walk-around. Throughout the field season, this

file will be backed up on a daily basis, even though the additional data collected each day is always inputted into the same file. Once validated through QA/QC procedures (SOP#16: Quality Assurance Project Plan), backups will be destroyed via deletion. As described in SOP #15: Post-Field Season, it is the responsibility of the GIS Specialist to complete the post-processing task of all GIS files created as part of this project. Once shapefiles have been processed, quality controlled by the GIS Specialist and Project Lead, posted to the project GIS folder, and uploaded to the GIS Server, the Project Lead should submit the Certification form to the Data Manager.

Training and Contact Information

Prior to implementing field work, a list of contact information for each person involved in the Lakes project will need to be submitted to the Network Data Manager. Contact information will include:

- First Name
- Last Name
- Middle Initial
- Organization
- Position Title
- Mailing Address
- Email Address
- Work Phone Number

Each person conducting field work as part of this protocol will need to follow the training procedures outlined in SOP #2: Field Crew Training. Log books developed in a standardized Excel format will need to be delivered to the Network Data Manager following the timeline listed above.

Reports

There is the potential for a variety of reports to be developed utilizing data collected as part of this lake monitoring project, including Annual Reports, Analysis and Synthesis reports, scientific publications, one page summary reports, and comprehensive reports.

Annual Reports and Analysis and Synthesis reports will be the responsibility of the Project Lead and should be submitted in the NPS Technical Report Series or Data Series format as appropriate, unless utilizing another series format for publication.

The reports described in SOP #19: Data Reporting and Analysis and any scientific publications should be submitted to the Network Data Manager upon completion.

Summary reports are one page reports completed by the Network staff and used to sum up information in the annual and analyses reports. These reports should be completed within 1 month of receiving the annual or analyses reports. These reports will be sent to all park employees and will provide links to the larger reports.

SOP #17: Data Transfer, Storage, and Archive (continued).

Lastly, the reports will be entered into NatureBib, the National Park Service natural resource bibliography application to catalog, search, and manage natural resource-related information sources pertaining to national parks. Additional information on NatureBib is available at: http://science.nature.nps.gov/im/apps/nrbib/. Reports and species information will be linked to NPSpecies, a master NPS database for documenting the occurrence and status of species in more than 270 national park units that contain significant natural resources. More information is available at: http://science.nature.nps.gov/im/apps/npspp/.

Data Storage

Project folders have been created for each monitoring protocol the KLMN plans to implement (Figure 1). Project folders contain five standard folders using a naming convention that includes the project title and one of the following: Documents, GIS, Data, Images, or Analysis. These five folders will contain all the data and information for a project as follows:

a) **Lakes_Documents.** This folder contains the reports, budgets, work plans, emails, protocols, contracts, datasheets, and agreements associated with a specific project.
b) **Lakes_GIS.** This folder contains shapefiles, coverages, layer files, geodatabases, GPS files, GIS/GPS associated metadata, and spatial imagery associated with a project.
c) **Lakes_Data.** This folder contains the KLMN Lakes database and .dbf files from the six field databases.
d) **Lakes_Images.** This folder contains any photographs related to the project and associated image metadata. In addition, copies of all photographs and metadata will be transferred into the KLMN Image database. Details on the KLMN Image database can be found in the KLMN Data Management Plan.
e) **Lakes_Analysis.** This folder will contain derived data and associated metadata created during analysis.

SOP #17: Data Transfer, Storage, and Archive (continued).

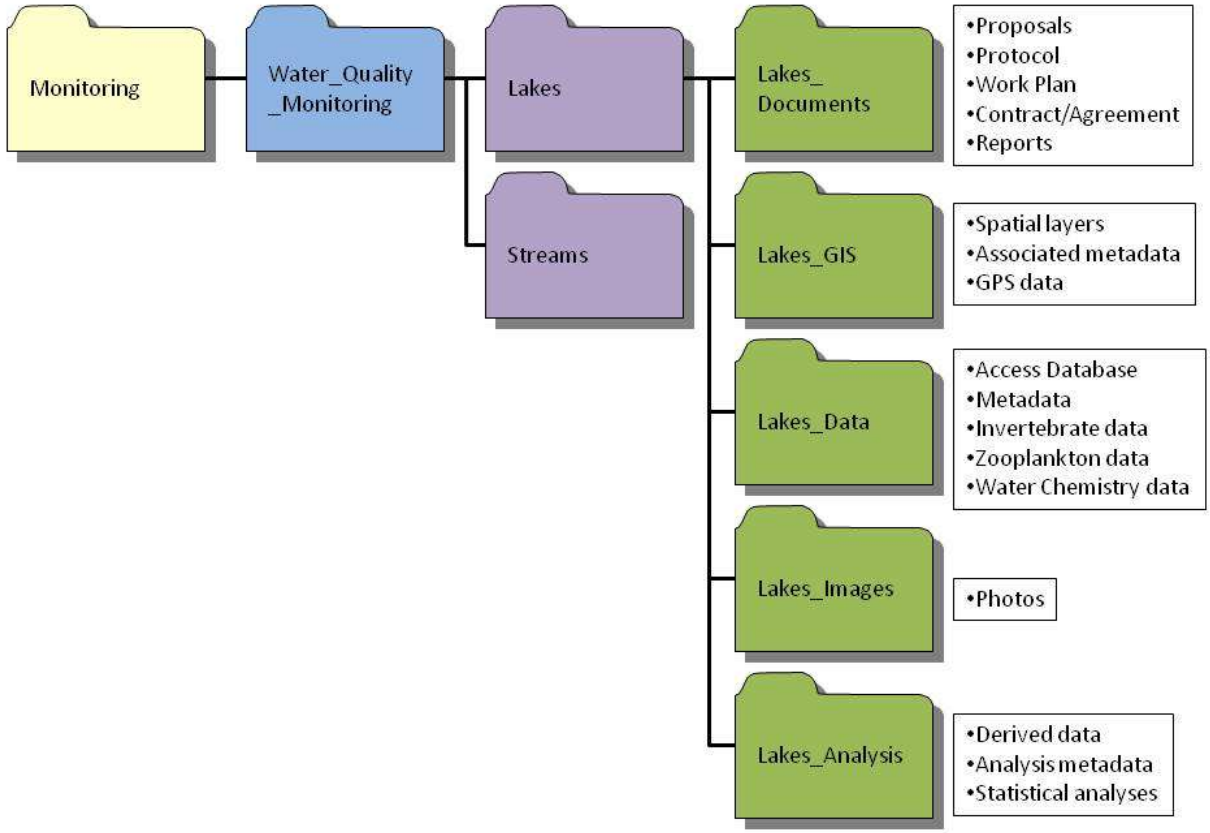

Figure 1. The Klamath Network Mountain Lakes and Ponds file structure to store all data and protocol information.

Network staff will have read-only access to all products to prevent changes to the information. If write access is needed, you will need to contact the Network Data Manager. It is the responsibility of all Network staff to inform the Data Manager when they have added new material to the project folder.

Storage, Backup, and Archiving

A copy of the project folder will be stored in the Network Archive drive whenever any new information is added to the folder. The Network Archive and Network drives are subject to all backup and archiving process described in the Network Data Management Plan. The Network relies on Southern Oregon University (SOU) for the backup and long-term storage requirements. Nightly backups are done by SOU to store information that has been edited. This is not a full backup but is intended to protect products that have been manipulated. This information is stored for a 1 week period before it is recycled. SOU begins a weekly full backup of their servers on every Friday and stores the files on tape drives. Backups are stored for 60 days before the tapes are reused. SOU will run quarterly backups on March 31st, June 30th, October 31st, and December

SOP #17: Data Transfer, Storage, and Archive (continued).

31st of each year. Files stored on a quarterly basis are maintained for 1 year before being recycled (Mohren 2007).

Despite the QA/QC measures in place, finding errors in datasets in the future is inevitable. The process for documenting the correction of such errors is detailed in SOP #16: Quality Assurance Project Plan. In such instances, archived data will not be corrected; however, an updated product will be placed into the archive drive along with the digital error and entry logs.

Table 1. Deliverable products, responsible individual, due data, and store location for all products developed while implementing the *Klamath Network Mountain Lakes and Ponds* Monitoring Protocol.

Deliverable Product	Primary Responsibility	Target Date	Instructions for Network
Contact Information	Project Lead	Two weeks prior to the start of the field season	Lakes Database Stored in the Lakes_Data Folder.
Metadata Interview Form	Project Lead	Prior to beginning the first field season and by Feb 1st of the following year when updates occur	Store in Lakes_Data[5], Use to create and revise full metadata.
Updated Data Dictionary	Project Lead	Prior to beginning the first field season and by Feb 1st of the following year when updates occur	Store in Lakes_Data[5], Use to create and revise full metadata.
Full Metadata (Parsed XML)	Network Data Manager	Prior to beginning the first field season and by March 1st of the following year	Store in Lakes_Data[5], Upload the Parsed XML Record to the NPS Data Store [2]
Protocol Changes (if made)	Project Lead	2 weeks prior to implementing the change	Store in Lakes_Document[5], Update Protocol on Websites and NPS Data Store, Send Copy to Parks
Data Certification Report	Project Lead	every time a product(s) is submitted	Store in Lakes_Document[5]
Field Data Forms, Fish Forms	Project Lead	Feb 1st of the following year	Scan Original, Marked-up Field Forms as PDF Files and Store in Lakes_Document[5]
Equipment, datasheet, event, and training logs	Project Lead	Feb 1st of the following year	Store in Lakes_Documents[5]
QA/QC Field Database	Project Lead	Prior to field crew being released	Store in Lakes_Data[5],
QA/QC Invert Database Upload	Project Lead	Feb 1st of the following year and prior to conducting analysis or summaries of reports	Store in Lakes_Data[5],
QA/QC Manta Database Upload	Project Lead	Feb 1st of the following year and prior to conducting analysis or summaries of reports	Store in Lakes_Data[5],

SOP #17: Data Transfer, Storage, and Archive (continued).

Table 1. Deliverable products, responsible individual, due data, and store location for all products developed while implementing the *Klamath Network Mountain Lakes and Ponds* Monitoring Protocol (continued).

Deliverable Product	Primary Responsibility	Target Date	Instructions for Network
Digital Photographs and Metadata	Project Lead	Feb 1st of the following year	Store in Lakes_Image[5], Copies of Photographs in Network Image Library, Copies of Image Metadata into Network Image Database linked to Photographs
Annual Report	Project Lead	June 1st of the following year and prior to starting any new fieldwork	Store in Lakes_Document[5], Upload to NPS Data Store[2], Send Copy to Parks, Post on the Network Internet and Intranet Websites
Analyses and Synthesis Report	Project Lead	Every three years on November 1st	
Other Publications	NPS Staff, Project Lead	As completed	
Other Records	Network Contact	Review for retention every Oct 1st	Digital Files that are Slated for Permanent Retention Should be Uploaded to the Network Lakes Project Folder. Retain or Dispose of Records Following NPS Director's Order #19 [4].

[1] The Network Image Library is a hierarchical digital filing system stored on the Network file servers. The image library is linked to an image database that stores metadata on each image.
[2] NPS Data Store is a clearinghouse for natural resource data and metadata (http://science.nature.nps.gov/nrdata). Only non-sensitive information is posted to NPS Data Store.
[3] NatureBib is the NPS bibliographic database (http://www.nature.nps.gov/nrbib/index.htm). This application has the capability of storing and providing public access to image data (e.g., PDF files) associated with each record.
[4] NPS Director's Order 19 provides a schedule indicating the amount of time that the various kinds of records should be retained. Available at: http://data2.itc.nps.gov/npspolicy/DOrders.cfm.
[5] The Network Lakes project folder located on the shared file server at the Network office. The project folder contains five folders including: Lakes_Documents, Lakes_Data, Lakes_Analysis, Lakes_GIS, and Lakes_Image used to separate and store data and information collected as part of the Klamath Network Mountain Lakes and Ponds monitoring.

Literature Cited

Mohren, S. R. 2007. Data management plan, Klamath Inventory and Monitoring Network. Natural Resource Report NPS/KLMN/NRR—2007/012. National Park Service, Fort Collins, Colorado.

SOP #17: Data Transfer, Storage, and Archive (continued).

KLMN Certification Form (also included in Appendix F)

1) Certification date: _____

2) Certified by: _____
Title: _____
Affiliation: _____

3) Agreement code: _____
Project title: _____

4) Range of dates for certified data: _____

5) Description of data being certified: _____
_____
_____
_____

6) List the parks covered in the certified data set, and provide any park-specific details about this certification.

Park	Details

7) This certification refers to data in accompanying files. Check all that apply and indicate file names (folder name for images) to the right:

_____ Hardcopy Datasheet(s): _____
_____ PDF Datasheet(s): _____
_____ Database(s): _____
_____ Spreadsheet(s): _____
_____ Spatial data theme(s): _____
_____ GPS file(s): _____
_____ Geodatabase file(s): _____
_____ Photograph(s): _____
_____ Data Logger(s) files: _____
_____ Other (specify): _____

SOP #17: Data Transfer, Storage, and Archive (continued).

_____ Certified data are already in the master version of a park, KLMN or NPS database. Please indicate the database system(s): _____

8) Is there any sensitive information in the certified data which may put resources at greater risk if released to the public (e.g., spotted owl nest sites, cave locations, rare plant locations)?

_____ No _____ Yes Details:

_____
_____
_____

9) Were all data processing and quality assurance measures outlined in the protocol followed?

Yes / No

If No, Explain _____
_____
_____

10) Who reviewed the products?
_____

11) Results and summary of quality assurance reviews, including details on steps taken to rectify problems encountered during data processing and quality reviews.
_____
_____
_____